

### EXAMINER'S AMENDMENT

1. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

Authorization for this examiner's amendment was given in a telephone interview with Kenneth Eiferman on 6/29/09.

2. Please replace Claims 1, 14, 17, and 22 with amended Claims 1, 14, 17, and 22 below.

1. (Currently Amended) A method for detecting similar objects in a collection of such objects, the method comprising:

processing, by a computer processor, a query to produce the collection of objects;

constructing a plurality of has tables collecting objects produced by processing the query; and, for each of two objects:

modifying a previous method for detecting similar objects so that memory requirements are reduced while avoiding false detections approximately as well as in the previous method, wherein the modifying comprises:

combining, by the computer processor, four samples of features into seven supersamples;

compressing each of the seven supersamples into sixteen bits of precision;

constructing ~~a plurality of~~ fifteen hash tables storing combinations of the supersamples;

comparing the supersamples using the fifteen hash tables; and

requiring a number of matching supersamples out of the seven supersamples in order to concluded that the two objects are sufficiently similar, wherein the number of matching supersamples is greater than a number of matching supersamples required in the previous method.

14. (Currently Amended) A method for determining groups of near-duplicate items in a search engine query result, the method comprising constructing a plurality of hash tables collecting items in the search engine query result and, for each of two items being compared:

combining, by a computer processor, four samples of features into each of seven supersamples;

compressing, by the computer processor, each supersample into 16 bits of precision;

constructing fifteen hash tables storing combinations of four supersamples;

using the fifteen hash tables to compare the supersamples; and

requiring five of the seven supersamples to match.

17. (Currently Amended) A computer-readable storage medium embodying machine instructions implementing a current method for detecting similar objects in a collection of such objects, wherein the current method comprises modification of a previous method for detecting similar objects so that memory requirements are reduced while avoiding false detections approximately as well as in the previous method, the current method comprising:

processing a query to produce the collection of objects;

constructing a plurality of hash tables ~~for the collection of~~ collecting objects produced by processing the query; and, for each of two objects,

combining four samples of features into each of seven supersamples, and

compressing each of the seven supersamples to sixteen bits of precision;

constructing ~~a plurality of fifteen~~ hash tables storing combinations of the supersamples;

comparing the supersamples using the fifteen hash tables; and

requiring a number of matching supersamples in order to conclude that the two objects are sufficiently similar, wherein the number of matching supersamples is greater than a number of matching supersamples required in the previous method.

22. (Currently Amended) A computer-readable storage medium embodying machine instructions implementing a method for determining groups of near-duplicate items in a search engine query result, the method comprising constructing a plurality of hash

tables collecting items in the search engine query result and, for each of two items being compared:

combining four samples of features into each of seven supersamples;  
compressing each supersample into 16 bits of precision;  
constructing fifteen hash tables storing combinations of four supersamples;  
using the fifteen hash tables to compare the supersamples; and  
requiring five of the seven supersamples to match.

***Allowable Subject Matter***

3. Claims 1-4, 8-10, 14-20, and 22 are allowed.
4. The following is an examiner's statement of reasons for allowance:

Multiple hash tables (some tables for objects produced by processing the query and 15 hash tables storing combinations of supersamples) being used for fingerprinting in detecting duplicate/similar objects combined with other claimed limitations are novel and non-obvious features of the present invention.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

**Conclusion**

5. Any prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Applicant is advised that, although not used in the rejections above, prior art cited on any PTO-892 form and not relied upon is considered materially relevant to the applicant's claimed invention and/or portions of the claimed invention.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brent S. Stace whose telephone number is 571-272-8372 and fax number is 571-273-8372. The examiner can normally be reached on M-F 9am-5:30pm EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Apu M. Mofiz can be reached on 571-272-4080. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

/Etienne P LeRoux/  
Primary Examiner, Art Unit 2161

/B.S./ (6/29/09)  
Examiner, Art Unit 2161